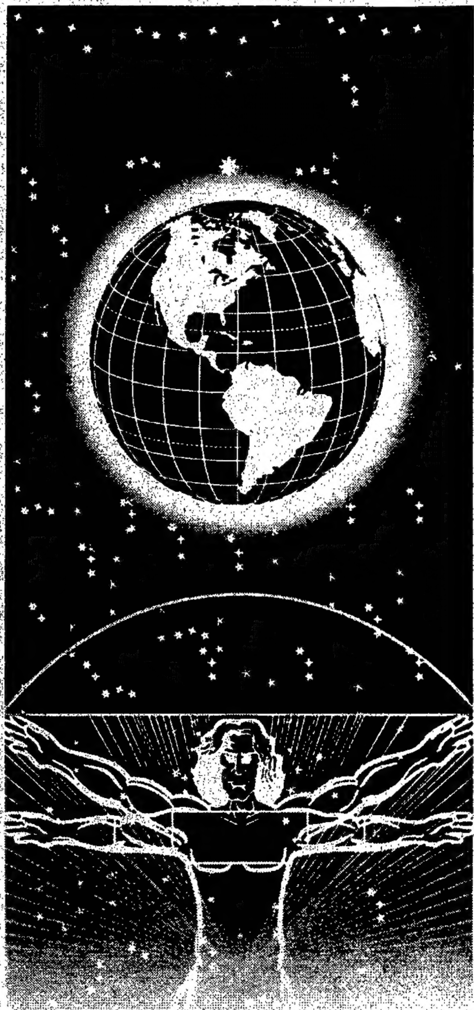
AL/AO-TR-1997-0025



UNITED STATES AIR FORCE
ARMSTRONG LABORATORY

"POPCORN" AS A TOOL FOR FUTURE COGNITIVE WORKLOAD ASSESSMENT: A CONCEPTUAL ANALYSIS

Shih-sung Wen

Universal Energy Systems, Inc.
4401 Dayton-Xenia Rd.
Dayton OH 45432

John C. Patterson

AEROSPACE MEDICINE DIRECTORATE
Clinical Sciences Division
Neuropsychiatry Branch
2507 Kennedy Circle
Brooks Air Force Base, TX 78235-5117

June 1997

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this technical report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication.

John C Patterson

JOHN C. PATTERSON, Ph.D.
Contract Monitor

K.F. Glifort, MD

KENNETH F. GLIFORT, Colonel, USAF, MC, CFS
Chief, Clinical Sciences Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
June 1997

3. REPORT TYPE AND DATES COVERED
Final Report

4. TITLE AND SUBTITLE
"POPCORN" as a Tool for Future Cognitive Workload Assessment: A Conceptual Analysis

5. FUNDING NUMBERS
C - F49620-85-C-0013
PE - 62202F
PR - 7755
TA - 26
WU - 47

6. AUTHOR(S)
Shih-sung Wen
John C. Patterson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Universal Energy Systems, Inc.
4401 Dayton-Xenia Road
Dayton, OH 45432

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Armstrong Laboratory
Aerospace Medicine Directorate
Clinical Sciences Division, Neuropsychiatry Branch
2507 Kennedy Circle
Brooks Air Force Base, TX 78235-5117

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AL/AO-TR-1997-0025

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

As weapon systems and their battlefield employment become increasingly complex, the understanding and assessment of the human operator becomes equally important. This project further developed a prototype cognitive workload assessment tool, "POPCORN." Starting with a single laboratory, single computer model, a PC based tool that is more widely usable was developed. In its current form, POPCORN contains multiple independent and dependent variables that can be easily manipulated at numerous levels from familiarization and training levels to virtually impossible. Further, a conceptual analysis was written based upon the authors' experiences with POPCORN as well as comparison of the tool with literature in the area. Based upon this analysis, we conclude that POPCORN is capable of measuring a wide range of cognitive processes relevant to human management of complex systems.

14. SUBJECT TERMS
Cognitive
Performance
Workload

15. NUMBER OF PAGES
20

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. Z39-18
298-102

CONTENTS

INTRODUCTION

Importance of Cognitive Workload Assessment

Workload assessment is of great interest to cognitive psychologists. Although "workload" has not been clearly defined, it consists of a set of task demands, operator efforts, and task performance (Gartner & Murphy, 1979). Cognitive psychologists try to understand how the intended task demands are perceived, processed, and performed by an aircraft operator.

Advancement in avionic technology has altered a great deal of the traditional man-machine relationship in the cockpit. The computer hardware and software in the aircraft can assume control and monitoring of a routine flight, handle repetitive tasks, and automatically diagnose, adjust, and correct mechanical errors. This achievement in automation has drastically reduced a large number of psychomotor or physical task demands (Hart & Sheridan, 1984). Cognitive demands, on the other hand, have become increasingly important for the pilot.

Today, computer-automation and its display systems constitute the key element of cockpit environment. The automated, sophisticated, high-speed man-made "brain" constantly interacts with the pilot and to receive commands under critical flight conditions. Thus, piloting is no longer a one-way communication from the pilot to the instrument. In most cases, it is neither a direct nor an analog control of the aircraft. The versatility of the artificial brain provides the pilot with instant messages for actions or reactions. It also lists various options for pilots to judge, choose, or reject in decision making. This constant feedforward and feedback between the human brain and the artificial brain inevitably means that cognitive processes, such as attention, perception, comprehension, reasoning, judgment, memory, learning, decision-making, and problem-solving, are dominant features in piloting.

Psychologists are increasingly aware of the types of task demands imposed on pilots. Questions they are trying to answer are: what cognitive processes are involved and what cognitive resources are available for a mission flight? The former is concerned with simulation of cognitive demands and the latter involves resources identification and utilization. Understanding the underlying cognitive processes of complex flight tasks would benefit pilot training by identifying and structuring priorities of tasks and specifying conditions under which optimal cognitive functions and outcomes could be achieved. Similarly, knowledge of a pilot's resources available for a mission flight could assist aviation engineers in designing a friendly cockpit environment with a feasible flight model which would best utilize human resources without causing overload (performance deterioration) or underload (boredom). Finally, a comprehensive, valid and efficient tool for cognitive evaluation would be useful for screening and clinical assessment. Hence, the importance of cognitive workload assessment can not be understated.

Approaches to Workload Assessment

According to Wierwille, et al (1979), workload assessment techniques could be evaluated by a matrix of two dimensions: workload methodology and operator behavior. Methods of workload assessment include subjective opinion (rating scale, interview, or questionnaire), spare

mental capacity (task analytic, secondary-task, occlusion), primary task measures (single or multiple measure, mathematical modeling), and physiological measures (single or combined measure). Berliner, et al (1964) categorized operator behavior into perceptual (searching for and receiving information, identifying objects, actions, events), mediational (information processing, problem-solving, and decision-making), communication, and motor (simple/discrete, complex/continuous).

Critical Problems Associated with Assessment Approaches

Most workload measures, in terms of measuring perceptual, cognitive, and communication workload, have had limited success (Casali and Wierwille, 1984; Hart and Sheridan, 1984; Hartman and McKenzie, 1979; McKenzie and Hartman, 1979; Rahimi and Wierwille, 1982; Wierwille, Williges, and Schiflett, 1979). The underlying causes of failure to validly assess or predict a pilot's task performance in reality can be attributed to: (a) the lacking of a consensus on the definition of workload and its performance standards (McKenzie and Hartman, 1979); (b) no theoretical foundation for the test development (Hart and Sheridan, 1984); (c) issues with sensitivity and intrusion in simulation procedures (Casali and Wierwille, 1984; Heffley, Clement, and Jewel, 1982); (d) inadequate attention to interactions between task loading and operator factors (O'Donnell, 1982), (e) irrelevance or meaningless in task assignment (Rabbitt, 1984; Witlin, 1984); and (f) a general dissociation between objective and subjective measures (Gopher and Braune, 1984; Vidulich and Wickins, 1983); further reducing validity, in our opinion, is the common use of static, single-mode tasks with no interactive, multitask, dynamic challenges.

A review, by this author, of task descriptions of some latest assessment batteries, such as Criterion Task Sets (CTS), Unified Tri-Service Cognitive Performance Assessment Battery (TS-CPAB), Basic Attributes Tests (BAT-Version 4), and other similar instruments, has revealed a number of psychometric flaws for their intended applications. First, most subscales or subcategory tasks were put together in a piecemeal fashion without clear rationale behind the arrangement. While each test has an empirical literature, the "battery" suffers from lack of validity studies. This was especially critical in dual-task or secondary-task assignments. Lack of Gestalt in task arrangement is distractive and could negatively affect the subject's willingness to actively participate in assessment and hence reduce the face validity of the instrument. Second, tasks were essentially measures of basic perceptual skills leaving many high-level cognitive processes untouched. Although cockpit facilities have been greatly reduced by introducing multifunctions display systems (Statler, 1984), many test developers remained committed to developing tasks of monitoring constancy or detecting changes on traditional panel displays. Thus, many high-level cognitive tasks, which the automation is unable to do (Hart and Sheridan, 1984), have been seriously ignored. Finally, the pilot's knowledge base, which affects an individual's perception and performance of task demands (Hart and Sheridan, 1984; Santilli, 1985), was omitted from current measures. Lack of, or inadequate, knowledge of specific task demands or unexpected task conditions could cause errors which might either lead to irreversible accidents or become an extraneous source of mental workload (Hart & Bortolucci, 1984).

Principles of an Effective Cognitive Workload Assessment

Flight is a purposive behavior. To validly measure cognitive workload in flight, the principles listed below should be followed.

A comprehensive flight model should be established to guide the development of assessment tasks. A comprehensive flight model should clearly define workload, specify hierarchical structures of man-machine relationship, describe conditions under which optimal results can be reached, and identify criteria against which behavior outcomes can be evaluated.

Lower-level cognitive processes, such as perceptual skills and attentional distributions, should be measured within a contextual framework of higher-level cognitive processes, such as metacognition, decision-making, or problem-solving. This will insure that all aspects of assessment are meaningful and relevant to the goal of flight and hence the result can be interpreted in a significant manner.

Task performance should reflect a human-environment dynamic system of interactions. Since cognitive processes in flight are constantly modified by results of interactions between the pilot and his environment, simulation tasks should allow the subject to see results of his actions and show records of his inputs.

Measurement instruments should be periodically revised in light of both technological advancement and strategic innovation. Recent development in workload assessment did reflect some degrees of concerns for technological advancement. However, the strategic innovation which is closely associated with the technological upgrade was very much unattended. What does automation mean to the pilot? How-does the pilot perceive enemy threats when the integrated fire/flight (IFF) system is on hand? Issues related to new aircombat strategies should be addressed in the development of assessment Instruments.

POPCORN as an Assessment Tool

A Supervisory Control Simulation program (POPCORN) was developed by NASA Ames Research Center (Hart, 1984). It was designed to simulate the cockpit environment of the future. The program, according its developer, will measure such cognitive processes as attention, monitoring, memory, rule learning and application, predicting, decision making, problem-solving, and automated system management.

OBJECTIVE OF THE RESEARCH EFFORT

The objective of this effort was to convert a supervisory control simulation system developed by NASA Ames Research Center to a microcomputer platform and to conceptually analyze the POPCORN as a cognitive workload assessment instrument by mapping its task performance into cognitive processes of piloting. This was an initial step toward validation of the instrument.

METHOD

An Examination of Cognitive Demands in Future Cockpit Environment

An Aerospace Medical Panel Symposium was held to address "Human Factors Considerations in High Performance Aircraft" (Hennessy, 1984). The symposium generally agreed that the future aircraft design and air combat tactics would impose many more cognitive demands than have traditional ones. The cognitive demands that were specifically mentioned for future flight were: selectivity of perception (attention), interpretation (comprehension), monitoring (supervisory control), anticipation, prediction, recognition and diagnosis of unexpected situations, learning and memory, inductive reasoning, judgment, rule application, risk, decision making, problem solving, and self-regulation.

Task Analysis of the POPCORN

Basic Features

The POPCORN consists of computer-generated performance tasks to simulate cognitive processes of future cockpit environment. According to its 1984 experimental version, the POPCORN includes five tasks and fifteen functions for each task. The five tasks are displayed horizontally at the bottom of the cathode-ray tube (CRT), and the fifteen functions are displayed vertically in top-down sequence on the right edge of the CRT. Each task is represented by a box which contains up to 20 events ("kernels"). When a task (e.g., Box*) and a function (e.g., "Open" a lid) are sequentially activated by a magnetic pen on a pad (with similar displays on the CRT), kernels will start "popping" out of the box and can be "consumed" (scored) by touching the PERFORM key at the bottom-right corner: otherwise, they will "die" (scored as error). The kernels in the activated box will begin to mill around with increasing speed the longer they remain in the box unperformed. Boxes can be simultaneously activated (linked) and can be stopped or delayed or changed. The experimenter can, by modifying task instructions or computer software, change the nature, speed, meanings, and values of tasks and functions, and length of trial to increase the flexibility, complexity, and demands of task performance.

Independent Variables

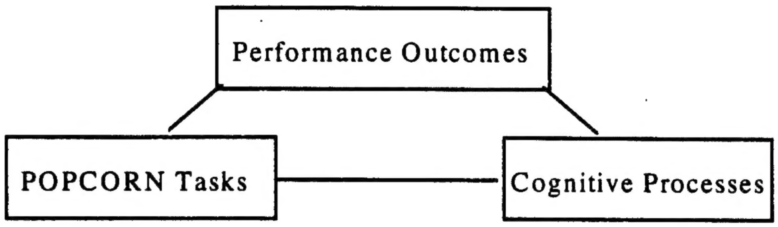
1. Number of tasks in action.
2. Number of functions used: OPEN (box), CLOSE (box), STUFF (elements), REMOVE (barriers from elements), STOP, CHANGE (colors), ADD (elements), SEED (elements), LINK (tasks), UNLINK (tasks), BACKUP (for diagnosis), CYCLE (door), MANUAL (for problem-solving).
3. Number of touches of PERFORMANCE key.
4. Task difficulty (Number of tasks X Number of functions).
5. Strategies (combinations of tasks and functions in various orders at different time).

Dependent variables

1. Score or penalty point (objective measure).
2. Satisfaction (subjective measure).
3. Stress (subjective measure).
4. Fatigue (subjective measure).

Mapping POPCORN Tasks into Cognitive Processes

Since POPCORN is a set of dynamic simulation tasks, it can evaluate interactions among tasks, cognitive (mediational) processes, and task performance outcomes. The following diagram depicts the dynamic relationship among variables of interest.



To proceed with the mapping, each task performance was conceptually simulated and sequentially checked with cognitive processes to find out the most appropriate matches in light of possible outcomes.

RESULTS

Cognitive Processes Measured by the POPCORN

Selective Attention

Attending to tasks selected, procedures to be followed, and outcomes to be expected.

Divided Attention

Attending to two or more tasks simultaneously activated.

Basic Perceptual Skills

Recognizing tasks to be completed, functions available for uses, and warnings of penalty; estimating speeds of popping kernels and possible scores can be earned; classifying tasks into various classes (completed vs. incomplete, active vs. inactive or delayed, regular vs. added, linked vs. unlinked); monitoring progress of each task; and predicting success against failure.

Judgment and Decision Making

Judging the quality and quantity of strategies used; determining underload, overload, or just-right workload; and deciding which options to choose to facilitate scoring, or avoid penalty.

Cognitive Strategies

Formulating or modifying task strategies to achieve maximum scores with high satisfaction but low stress and fatigue.

Learning

Acquiring and applying tasks-related rules; and modifying behavior through feedback.

Memory

Sensory store for each popping kernel; short-term memory for each action taken or score earned; and long-term memory of strategies used, rules followed, problems solved, mistakes made, and total scores accumulated.

Problem Solving

Solving the built-in problems in tasks by following predetermined procedures.

Metacognition

Knowing what one wants to do, what one has done, how tasks are to be done, and how many mental resources are still available; and knowing what one knows and what one does not know in dealing with tasks.

In summary, as the result of this conceptual analysis, the POPCORN appears to be able to meet the need of assessing major cognitive demands of future cockpit environment, although some are more adequately measured than others.

Unique Characteristics of the POPCORN

The impressive capacity of the POPCORN to measure a broad range of cognitive processes could be attributed to, among other things, its three unique features.

Man-Machine Interaction

From the beginning to the end, the whole simulation program is a process of action-and-interaction between the operator and the task. Perceptions, efforts, and outcomes are changing as time and tasks progress. It considers the operator as an active, purposive, and adaptive information processing system as did Newell and Simon (1972). As a result, all efforts, decision-making, strategies and other cognitive processes are "naturally" imbedded in the seemingly non-intrusive, game-like task performance. Consequently, human interest can be maintained; the process of task attainments can be analyzed; and the outcome can be meaningfully interpreted.

Metacognition Assessment

One advantage of having a complete, meaningful, and interactive simulation system is that it can be used to measure the pilot's metacognition. Metacognition is a superordinate executive system which overviews and guides on-going cognitive activities (Meichenbaum, 1977). For a pilot, knowing what he knows, what he is doing, and how things can be done is essential in determining courses of action. Perhaps, this tool can decide how much a pilot will benefit from

his own experience. Further, POPCORN may provide salient information on cognitive efficiency or superior cognitive performance and functioning.

Future Assessment Tool

The POPCORN was designed with considerations of high degree of automation and high demands of cognitive processes in future aircraft. In the cockpit, automation lets the pilot "fall behind;" he has to manage to "stay ahead" (Richter, 1984). POPCORN provides a superior test bed to understand cognitive processing in aviation.

SUMMARY AND CONCLUSION

The purpose of this project has been to convert a cognitive workload task (POPCORN) for use with MS-DOS-based microcomputers and to evaluate its utility as a cognitive assessment tool. The programming has been completed and, while certain cosmetic improvements could be made, it is now a useable tool for further research.

Although many current workload assessment instruments have had some degree of success, they are characterized by a number of problems: a consensus on the definition of workload is still lacking; a comprehensive theoretical foundation for simulation is not available; intrusion effects on measurement sensitivity are unsolved; man-machine interactions are ignored; task assignments are mostly irrelevant or meaningless; discrete molecular tasks are put together in a piecemeal manner to simulate the continuous and complex molar task of flying; few high-level cognitive skills are being measured; and the knowledge base of piloting is not properly included.

An analysis of the cognitive demands of future aircraft and combat strategies indicates that a new workload measurement instrument is needed to adequately assess high-level cognitive processes of piloting. Further analysis of a recently developed workload assessment tool, POPCORN, shows that it can monitor selective and divided attention, basic perceptual skills, judging and decision-making, cognitive strategies, learning, memory, problem solving and metacognition. The study concluded that the POPCORN is an interactive simulation system, capable of measuring metacognition, and that it is a unique example of assessment tools of the future.

RECOMMENDATION

Validity is the major quality with which a measurement can be properly evaluated. Low validity and poor generalizability of current workload assessment instruments have prompted two possible options: continuously modifying current instruments or constructing a new-generation assessment tool. The Basic Attributes Tests (BAT) and Unified Tri-Service Cognitive Performance Assessment Battery (CPAB) and other similar instruments are the result of the former attempt and the POPCORN the latter. On the ground specified in the present analysis, POPCORN should be further considered as a cognitive workload assessment tool of the future.

The result of this study suggests that the POPCORN is capable of measuring a large number of important cognitive processes of piloting. This study is only a first step toward validation of the instrument; a long-term empirical validity study should follow. To achieve this goal, however, a short-term empirical study is recommended to systematically investigate how a selected cognitive process is simulated by POPCORN tasks.

ACKNOWLEDGMENTS

This research was supported by the Air Force Office of Scientific Research, through the Summer Faculty Research Program which was conducted by Universal Energy System under Contract F49620-85-C-0013.

Thanks are due to Dr. Bryce Hartman for his support and encouragement, to Mrs Rosie M. Anguiano and Mrs Julie Yochum for typing the report.

REFERENCES

Berliner, C., Angell, D., Shearer, D.J. (1964). Behaviors, measures, and instruments for performance evaluation in simulated environment. Paper presented at the Symposium and Workshop on the Quantification of Human Performance, Albuquerque, New Mexico, August, 1964.

Casali, J.G., Wierwille, W.W. (1984). On the measurement of pilot perceptual workload: A comparison of assessment technique addressing sensitivity and intrusion issues. Ergonomics, 27, 1033-1050.

Gartner, W.B., Murphy, M.R. (1979). Contents of workload. In B.O. Hartman & R.E. McKenzie (Eds), Survey of methods to assess workload. AGARD-AG-246.

Gopher, D., Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? Human Factors, 26, 519-532

Hart, S.G. (1984). POPCORN: Subtask definition for a supervisory control simulation. NASA Ames Research Center, Moffett Field, CA.

Hart, S.G., & Bortolussi, M.R. (1984). Pilot errors are a source of workload. Human Factors, 26, 545-556.

Hart, S.G., Sheridan, T.B. (1984). Pilot workload, performance, and aircraft control automation. In R. T. Hennessy (Ed), Human factors considerations in high performance aircraft, AGARD-CP-371.

Hartman, B.O., McKenzie, R.E. (1979). Survey of methods to assess workload, AGARD-AG-246.

Heffley, R.K., Clement, W. F., Jewell, W.F. (1982). Overview of work in progress on non-intrusive assessment of pilot workload and pilot dynamics. Proceedings of the workshop on flight testing to identify pilot workload and pilot dynamics. AD-129-333.

Hennessy, R. T. (1984). Human factor considerations in high performance aircraft. AGARD-CP-371.

McKenzie, R.E., Hartman, B.O. (1979). Some insights relative to the man-machine system: An overview of ten years of research. In B.O. Hartman & -E. McKenzie (Eds), Survey of methods to assess workload. AGARD-AG-246.

Meichenbaum, D. (1977). Cognitive-behavior modification: An integrative approach. New York: Plenum Press.

Newell, A., Simon, H.A. (1972). Human problem solving. Englewood Cliffs, N.J.: Prentice-Hall.

O'Donnell, R.D. (1982). Historical foundations of the AFAMRL workload program. Proceedings of the workshop on flight testing to identify pilot workload and pilot dynamics, 53-67.

Rabbitt, P. (1984). The control of attention in visual search. In R. Parasuraman & D.R. Davies (Eds), Varieties of attention, New York: Academic Press.

Rahini, M., Wierwille, W.W. (1982). Evaluation of the sensitivity and and intrusion of workload estimation techniques in piloting tasks emphasizing mediational activity. Proceeding of the 1982 international conference on cybernetics and society, 593-597.

Richter, K. (1984). Impact of future aircraft characteristics on pilot performance and cockpit design. In R.T. Hennessy (Ed), Human factors considerations in high performance aircraft. AGARD-CP-371.

Santilli, S.R. (1984). Information processing and you. Tac Attack, June, 6-11.